**EXPLORATORY DATA ANALYSIS MINI PROJECT 2**

*Authors: Vikrant Deshpande, Tanvi Kolhatkar, Saishree Godbole*

# INTRODUCTION

In this project, we look at a dataset of home selling-prices in King County, Washington, between May 2014 and May 2015. The dataset contains 21,613 homes and 21 descriptive features. We will build a multiple regression model with the following variables: y is the outcome variable, and there are two explanatory variables:

1. A numerical explanatory variable x1: sqft living, which refers to the size of a house in square feet of living space. It is worth noting that 1 square foot equals 0.09 square meters.

2. A categorical explanatory variable x2: housing condition, a five-level categorical variable with 1 indicating "poor" and 5 indicating "outstanding."

3. a. Check if the relationship can be fitted by a linear model and explain any differences.

b. Analyze the trends of life expectancy over time for individual continents and respective countries.

c. Check if any other factors affect the house-price apart from the Square Footage, Condition, Bathrooms,and Grade.

eda